

ADVANCED TECHNOLOGY DEVELOPMENT FOR IMAGE GATHERING, CODING, AND PROCESSING

Friedrich O. Huck
NASA Langley Research Center

SUMMARY

Our research activities consist of three overlapping areas. 1) Information theory and optimal filtering are extended rigorously to visual information acquisition and processing. The goal is to provide a comprehensive methodology for quantitatively assessing the end-to-end performance of image gathering, coding, and processing. Information theory allows us to establish upper limits on the visual information that can be acquired within given constraints. Optimal filtering allows us to establish upper limits on the performance that can be attained for specific tasks, even if these tasks require adaptive or interactive processing. 2) Focal-plane processing techniques and technology are developed to combine effectively image gathering with coding. The emphasis is on low-level vision processing akin to the retinal processing in human vision. Our approach includes the familiar lateral inhibition, the new intensity-dependent spatial summation, and parallel sensing/processing networks. 3) A breadboard adaptive image-coding system is being assembled. This system will be used to develop and evaluate a number of advanced image-coding technologies and techniques as well as to research the concept of adaptive image coding. The idea of adaptive image coding is to use knowledge-based supervision to autonomously select the appropriate data compression scheme based on (a) the properties of the acquired image data (and changes in these properties), (b) the visual information required by the investigator, and (c) the available channel capacity. The system would have continual access to the investigators' knowledge of what is significant. Under the investigators' supervision, it could adaptively select image-coding and feature-classification schemes based on properties of the target such as spatial structure, texture, and spectral reflectance.

1. INFORMATION THEORY AND OPTIMAL FILTERING

The performance of (digital) image-gathering systems is constrained by the spatial-frequency response of optical apertures, the sampling passband of photon-detection mechanisms, and the noise generated by photon detection and analog-to-digital conversion. Biophysical limitations have imposed similar constraints on natural vision. Visual information is inevitably lost in both image gathering and low-level vision by aliasing, blurring, and noise. Therefore, it is no longer permissible to assume sufficient sampling as Shannon and Wiener could do in their classical works, respectively, on communication theory and optimal filtering for time-varying signals. Nevertheless, the digital-processing algorithms (for image restoration, edge enhancement, etc.) found in the currently prevailing literature assume sufficient sampling, whereas image-gathering systems are ordinarily designed to permit considerable insufficient sampling. This fundamental difference between assumption and reality has caused unnecessary limitations in the performance of digital image gathering, coding, and processing. It also has led to unreliable conclusions about the correct design of image-gathering systems for visual information processing (as opposed to image reconstruction without processing, e.g., commercial TV) and about the actual performance of image-coding schemes for tasks which involve digital image processing.

Our analyses so far have shown that the combined process of image gathering and optimal processing (see fig. 1) can be treated as a communication channel if (and only if) the image-gathering degradations are correctly accounted for. Correctly restored images gain significantly in fidelity (similarity to target), resolution (minimum discernible detail), sharpness (contrast between large areas), and clarity (absence of visible artifacts). These improvements in visual quality are obtained solely by the correct end-to-end optimization without increase in either data transmission or processing. These results have encouraged us to extend our analyses to various image coding schemes and the associated image-restoration and feature-extraction algorithms.

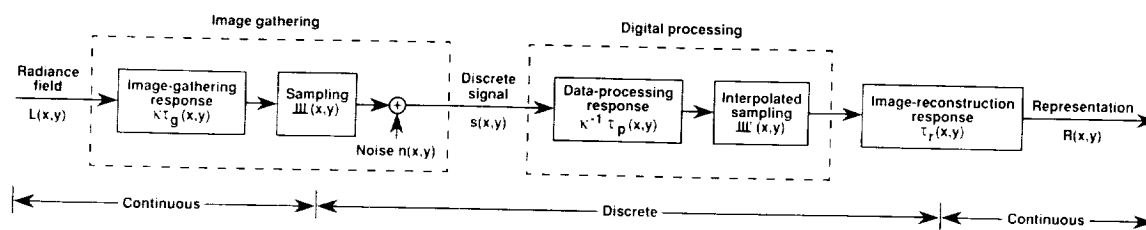


Figure 1. Model of image gathering, digital processing, and image reconstruction.

This work is summarized in references 1 through 6.

1. F. O. Huck, C. L. Fales, D. J. Jobson, S. K. Park, and R. W. Samms, "Image-plane processing of visual information," *Applied Optics*, Vol. 23, No. 18, pp. 3160-3167, 1984.
2. F. O. Huck, C. L. Fales, N. Halyo, R. W. Samms, and K. Stacy, "Image gathering and processing: Information and fidelity," *J. Opt. Soc. Am. A2*, 1644-1666 (1985).
3. F. O. Huck, C. L. Fales, J. A. McCormick, and S. K. Park, "Image-gathering system design for information and fidelity," *J. Opt. Soc. Am. A5*, 285-299 (1988).
4. C. L. Fales, F. O. Huck, J. A. McCormick, and S. K. Park, "Wiener restoration of sampled image data: End-to-end analysis," *J. Opt. Soc. Am. A5*, 300-314 (1988).
5. J. A. McCormick, R. Alter-Gartenberg, F. O. Huck, "Image gathering and restoration: Information and visual quality," *J. Opt. Soc. Am. A6*, July (1989).
6. F. O. Huck, S. John, J. A. McCormick, and R. Narayanswamy, "Image gathering, coding, and restoration: Information efficiency and visual quality." *Visual Information Processing*, Williamsburg, VA, 10-12 May, 1989.

2. FOCAL-PLANE PROCESSING TECHNIQUES AND TECHNOLOGY

Image gathering and coding are commonly treated as tasks separate from each other and from the digital processing used to restore and enhance images or extract features such as primal sketches or contour outlines. However, if we implement the edge enhancement required to draw primal sketches with focal-plane processing by properly combining optical response with lateral inhibition (as depicted by figs. 2 and 3), then a number of advantages can be gained. These advantages include improved resolution by a factor of four, improved robustness to noise, reduced data processing by two orders of magnitude, and reduced data transmission by nearly one order of magnitude. The implementation of a sensor-array focal-plane processor with lateral inhibition is supported by an SBIR contract (NAS1-18287, Phase II). If this approach could be extended to include not only edge enhancement but also edge detection so that only significant primal sketches need to be transmitted, then it would be possible to reduce data transmission by two orders of magnitude without loss of the improved accuracy in edge location.

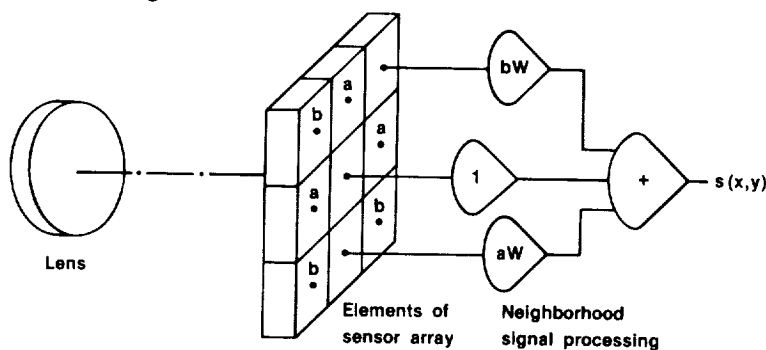


Figure 2. Model of optics and sensor array with focal-plane processing.

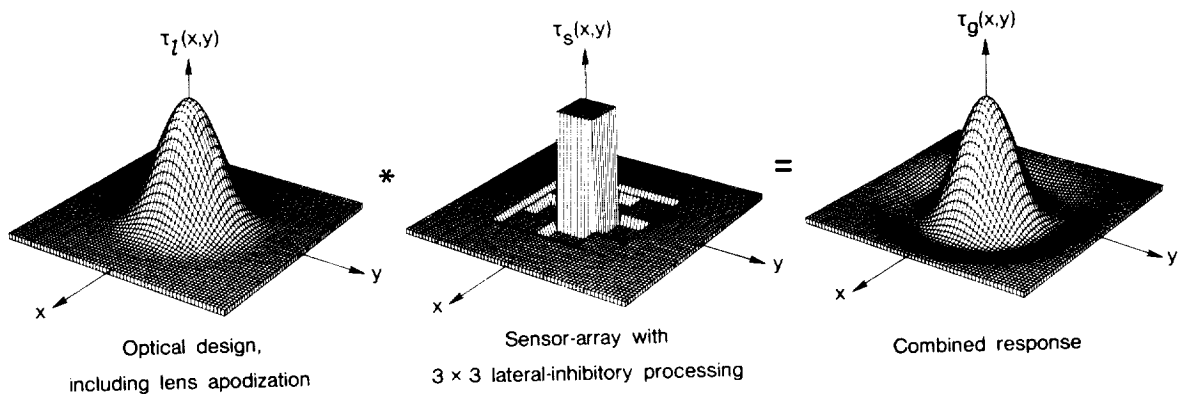


Figure 3. Laplacian-of-Gaussian response created by optical design and lateral-inhibitory processing.

Our major emphasis is on the development of an image-coding method that we refer to as local intensity adaptive image coding. This work is supported by two SBIR contracts (NAS1-18664 and NAS1-18850, both Phase II).

Local intensity adaptive image coding consists of an innovative model of processing in the human retina referred to as Intensity Dependent Spread (IDS) and some additional logic to extract contour outlines and reflectance ratios at the boundary of two surfaces. Figure 4 is a schematic representation of the IDS model. The line of detectors represents a slice through a two-dimensional array of detectors. When an optical image or light distribution falls on the detector array, then each detector sends a signal into a network, where it spreads out. Each channel, in turn, sends out a signal that is the sum of all the signals that arrive in its location in the summation network. The special property of the IDS model has to do with the way the signal from each detector spreads in the summation network. As depicted in the lower half of figure 4, the magnitude of the signal at its center is proportional to the intensity of the light falling on the detector, and the spread of the signal is inversely proportional to this intensity. The total volume under the spread remains constant. That is all there is to the model. It has been demonstrated, in reference 8, that this simple space-variant model of image processing has many of the properties of human visual perception.

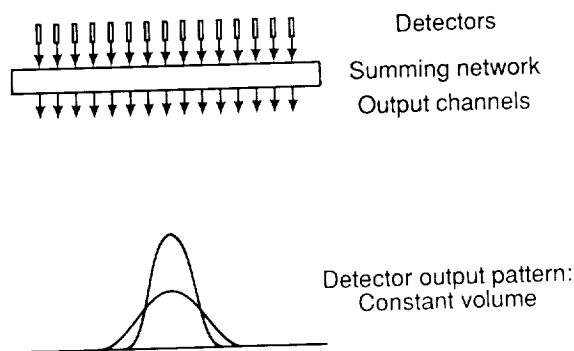


Figure 4. Schematic representation of the IDS model.

Figure 5 shows the response of the IDS processor to a spot, or point source, that is brighter than the background and to a step-type edge. Each detector spreads its signal as governed by the intensity of the light falling on the detector. For example, all of the spreads for the uniform background in figure 5(a) are the same except for the one detector that is more brightly lighted. Its spread is higher and narrower.

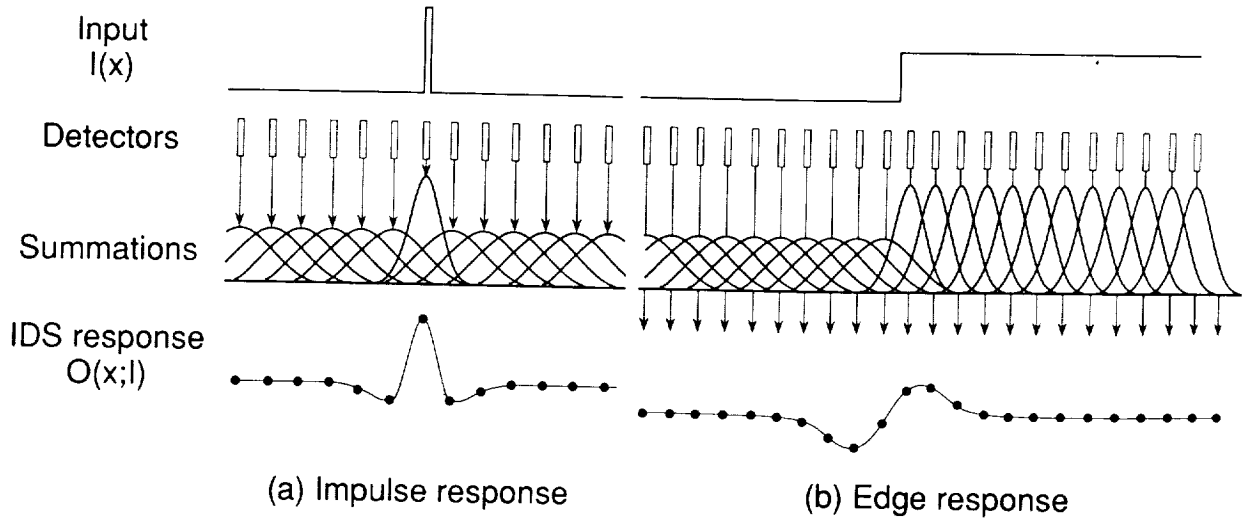


Figure 5. The IDS response to a point source and a step-type edge.

Each output channel just adds up all of the contributions it receives. The result of this processing is shown as the output signal. As can be seen, the IDS response to a point source has a similar shape as the response of Marr and Hildreth's familiar Laplacian of Gaussian ($\nabla^2 G$) operator for enhancing edges (see fig. 2). In fact, the IDS processor exhibits center-surround antagonism and all other manifestations of bandpass filtering that have made the $\nabla^2 G$ operator a favorite algorithm for low-level vision processing. However, the IDS response is nonnegative and spatially variant. As we will show in the next three figures, the IDS processor accounts for several familiar perceptual phenomena of human vision that make it a highly robust low-level vision operator.

First let us compare the IDS operation to conventional imaging. Figure 6 shows intensity profiles taken across conventional and IDS images of a step-type edge input for three illuminations. Conventional image-gathering yields a blurred representation that is visually representative of the target if the signal-to-noise ratio (SNR) is sufficiently high. As the illumination decreases, the representation gets buried in the noise. Image gathering with the IDS processor yields a target representation that consists of pulses. The one-crossing of each pulse locates the position of an edge in the target. The peak and trough values of the pulse are proportional to the ratio of the reflectances at the two sides of the edge, entirely independent of illumination. As the illumination decreases, the width of the pulse becomes broader (thereby trading resolution for sensitivity), but the accuracy of the one-crossing is unimpaired. For machine vision, this property means that edge detection for determining structure is highly robust in widely variant illumination.

Next let us compare the IDS operation to edge detection with the linear $\nabla^2 G$ operator as well as to conventional imaging. Figure 7 shows intensity profiles taken across conventional images and outputs from the $\nabla^2 G$ and IDS operators for two illuminations, high and low. Noise is disregarded for simplicity. The peak and trough values of the $\nabla^2 G$ pulses are proportional to both illumination and reflectance. Therefore, it is not possible to characterize the reflectance properties of the target independent of illumination. However, the peak and trough values of the IDS pulses are proportional only to the reflectance changes. This striking property of the IDS processor mimics human visual perception (Weber's law) and makes it possible to extract the reflectance ratio at the boundary of two areas.

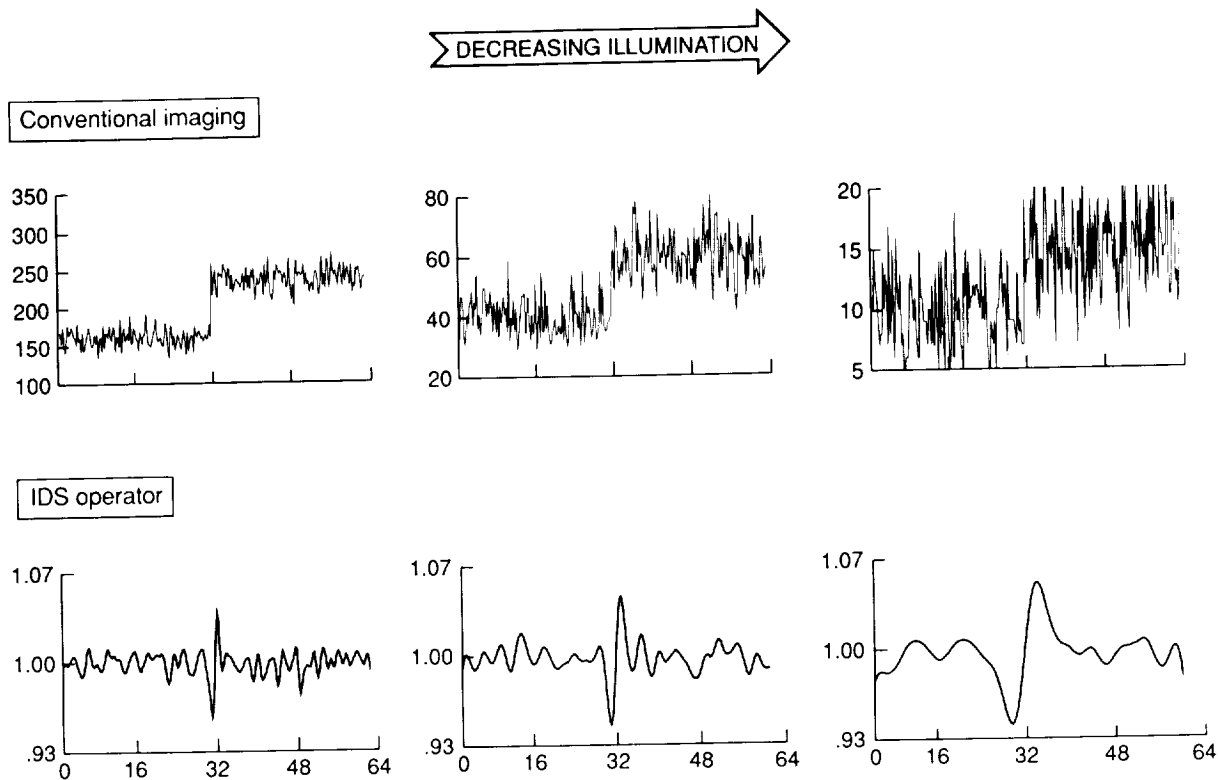


Figure 6. Comparison of IDS operation to conventional imaging.

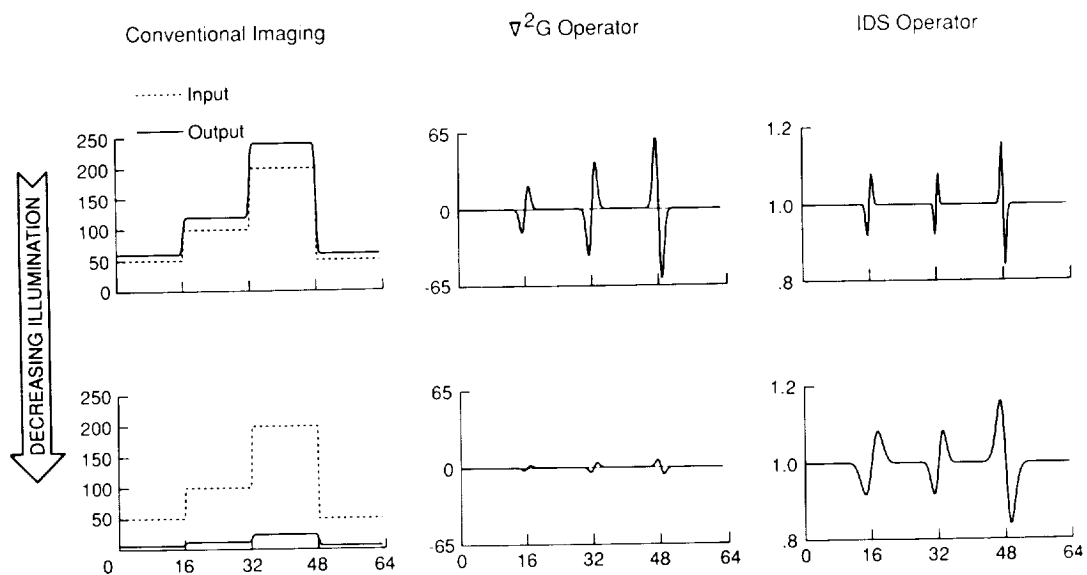


Figure 7. Comparison of IDS operation to conventional imaging and $\nabla^2 G$ operation.

Figure 8 illustrates a particularly important characteristic of the IDS filter, namely, the robustness of its reflectance representation to local variation in illumination (e.g., shadow). The recovered target 8(d) resembles the original one 8(a) and not the shadowy one 8(b), which is the one that was filtered. Traces of the shadow degradation can be seen in the modest loss of accuracy in the

actual transition as the illumination decreases. An important extension of IDS processing would be to extract color. Color then could be correctly detected independent of illumination.

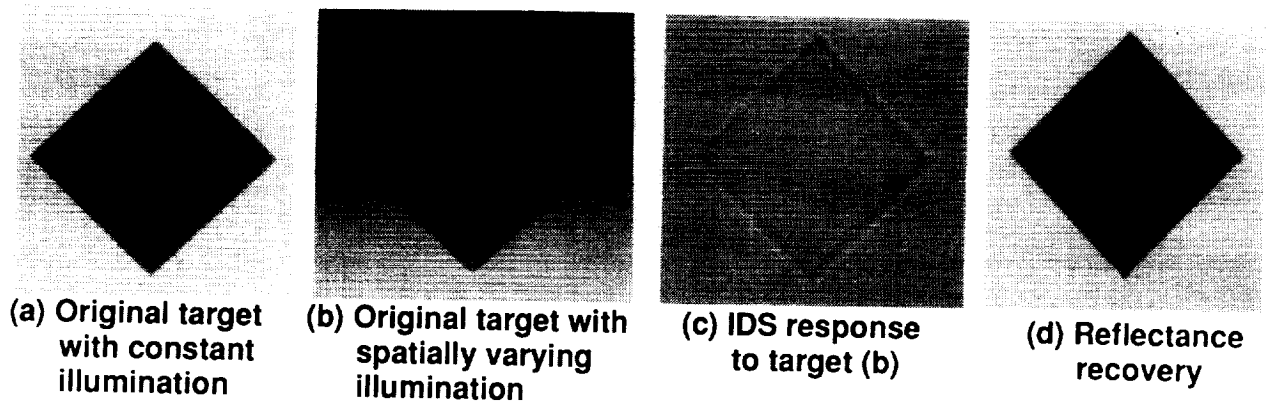


Figure 8. IDS response to target with spatially varying illumination.

Odetics, Inc., is now under contract with LaRC (NAS1-18664, Phase II) to develop a hardware implementation of the IDS processor (see fig. 9). This processor will be capable of handling image data at real-time TV rates (30 frames per second). It will be implemented on several boards for the DATACUBE of Sun image-processing work stations. These boards are expected to become commercially available in Fall 1989.

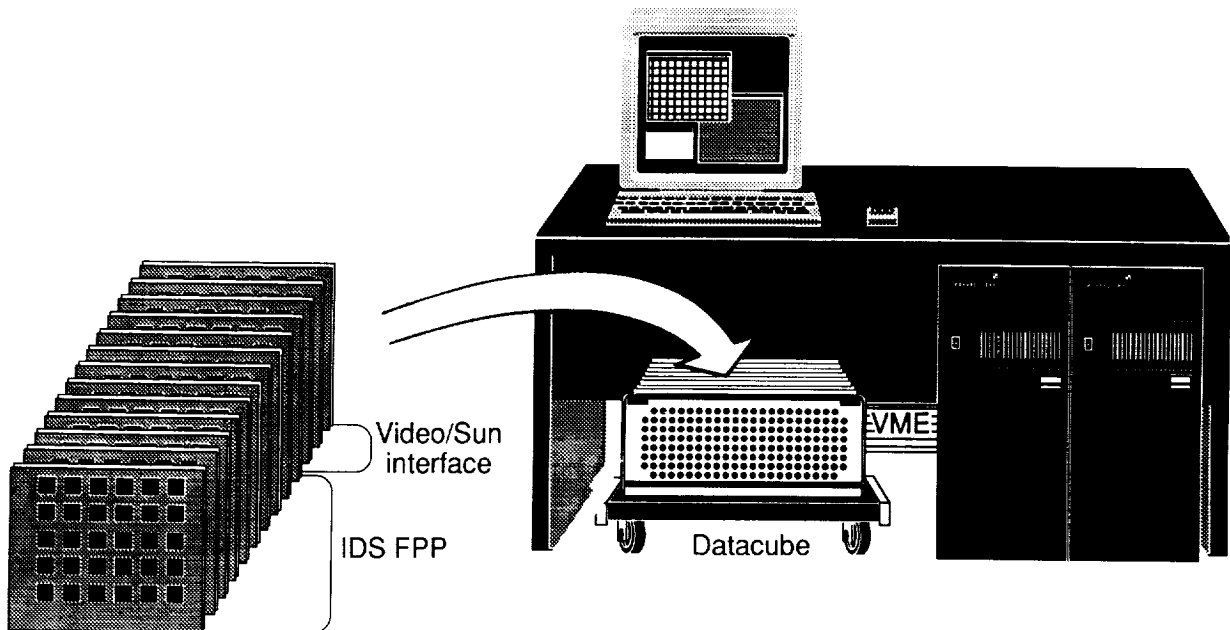


Figure 9. Implementation of IDS for real-time operation in Sun workstation.

The full potential of the IDS processor for data compression as well as image enhancement and feature extraction is realized, of course, only when it is implemented as a focal-plane processor, or "retinal camera," depicted in figure 10. The present design of the IDS processor for Sun workstations could be implemented in one 5" by 5" board with 8 VLSI chips. A more advanced approach is the parallel asynchronous focal-plane image processor depicted in figure 11. This processor, which is being developed under a contract with LaRC (NAS1-18850, Phase II), is representative of a new class of devices that permit full two-dimensional parallel readout and processing perpendicular to the focal plane. Advantages over conventional image-gathering and processing techniques include rapid parallel distributed processing, high dynamic range, and the elimination of conventional charge

transfer, multiplexing, and preamplifiers. Vision processing potentially could be performed several orders of magnitude faster than with conventional approaches. Moreover, parallel processing would be ideal for tasks like visual pattern recognition. However, the development of this approach is still in its initial experimental stage.

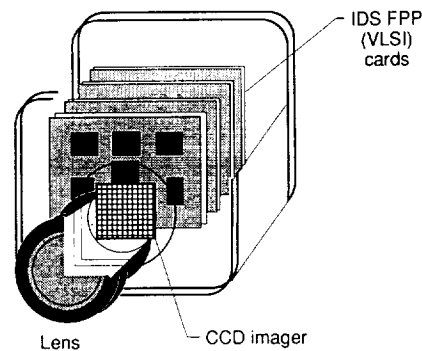


Figure 10. IDS focal-plane processing camera.

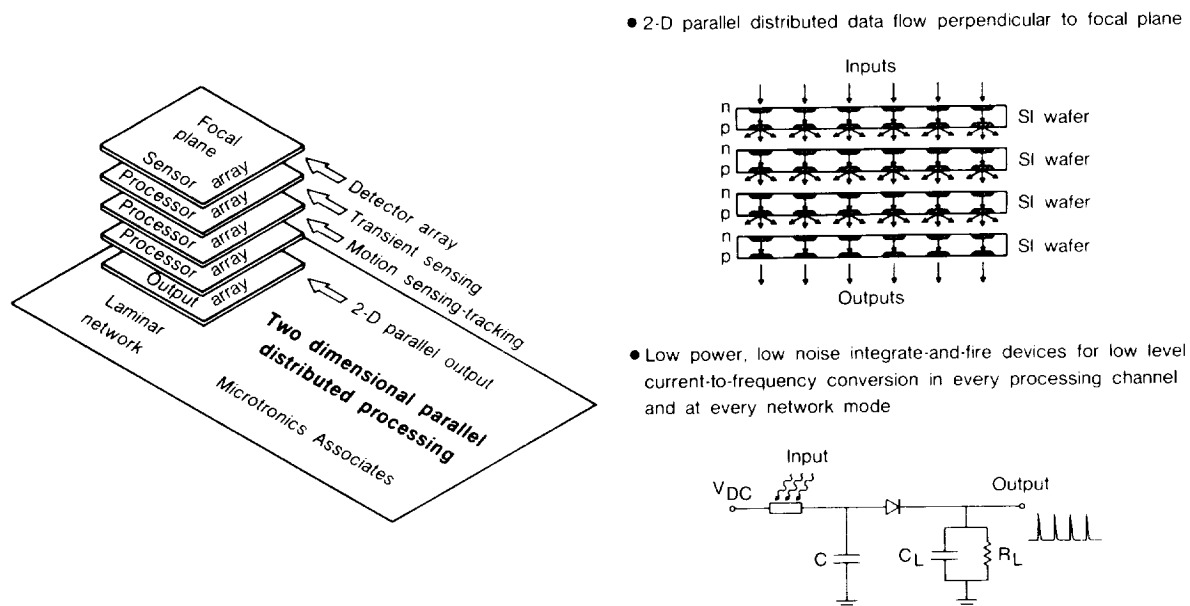


Figure 11. Parallel asynchronous focal-plane image processor.

This work is summarized in references 7 through 12.

7. R. Alter-Gartenberg, C. L. Fales, F. O. Huck, and J. A. McCormick: "End-to-end performance of image gathering and processing for edge detection," Computer Vision, Graphics and Image Processing, submitted.
8. T. N. Cornsweet and J. I. Yellott, Jr., "Intensity-dependent spatial summation," J. Opt. Soc. Am. 2, pp. 1769-86 (1985).
9. R. Alter-Gartenberg and F. O. Huck, "Feature extraction from intensity-dependent spatial summation," J. Opt. Soc. Am. A, submitted.
10. R. Alter-Gartenberg, R. Narayanswamy, and K. S. Nolker, "From primal sketches to the recovery of intensity and reflectance representations," Visual Information Processing, Williamsburg, VA, 10-12 May, 1989.
11. D. D. Coon and A. G. U. Perera, "Integrate-and-fire coding in Hodge-Huxley circuits employing silicon diodes," Solid State Electronics 31, 851, 1988.
12. D. D. Coon and A. G. U. Perera, "Integrate-and-fire dynamics and spike train information coding in neuron equivalent circuits employing silicon diodes," Neural Networks 1, 1988.

3. ADAPTIVE IMAGE CODING SYSTEM

The problem of image gathering and data compression for the high-resolution, high-frame-rate video system is such that no single data-compression method, or even just a few methods, can satisfy most of the diverse requirements. For some experiments the data-compression requirements even change with time. Consequently, we have started to investigate the concept of adaptive image coding. This investigation is supported by one SBIR contract (Phase I). We are currently seeking additional SBIR contracts under the subtopic 07.01 entitled Focal Plane Image Processing.

Figure 12 presents a block diagram of the basic adaptive image coding system. The idea of this system is to use knowledge-based supervision to autonomously select the appropriate data compression scheme based on (a) the properties of the acquired image data (and changes in these properties), (b) the visual information required by the investigator, and (c) the available channel capacity. The system would have continual access to the investigators' knowledge of what is significant. Under the investigators' supervision, it could adaptively select image-coding and feature-classification schemes based on properties of the scene such as spatial structure, texture, and spectral reflectance.

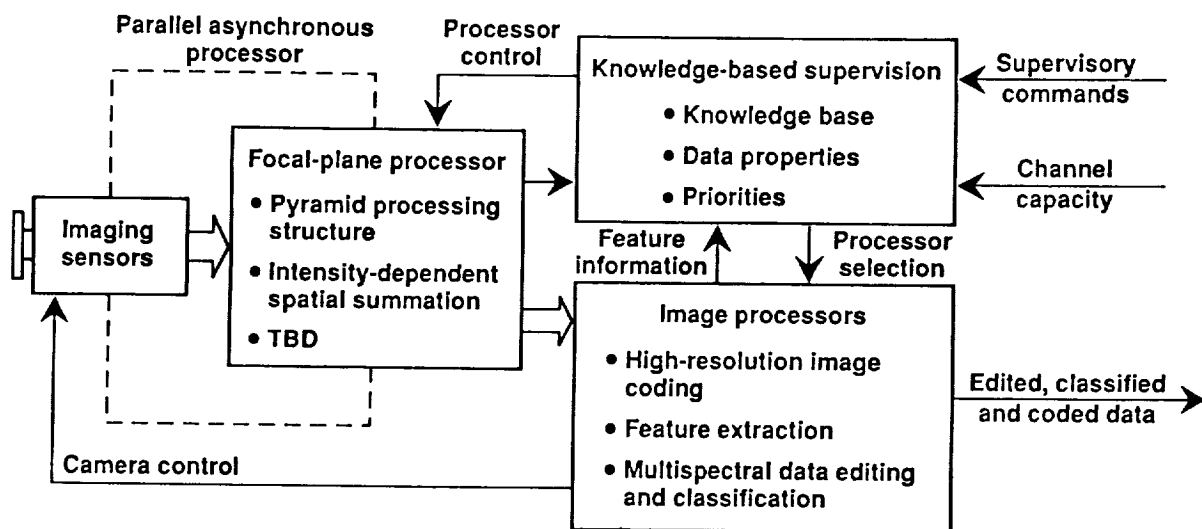


Figure 12. Schematic depiction of adaptive image coding system.

For example, consider figures 13 and 14. In one extreme, if the target of interest is well illuminated, as it is for figure 13, and the goal is to reconstruct an image of this target with pleasing visual quality, then the appropriate data-compression scheme should allow only those degradations to occur that are benign with regard to visual quality. Furthermore, Optivision's OPTIPAC coding system, for which these results are shown, allows the relationship between visual quality and amount of data compression to be continuously adjusted. This adjustment, for example, could be autonomously controlled by the channel capacity that is available for transmitting the desired images.



Figure 13. Data compression for well illuminated target. The compression ratio varies from CR = 22:1 to 92:1 for the original color pictures shown here only in black and white. (Courtesy of Optivision, Inc.)

In the other extreme, consider the scene shown in figure 14. An extreme range of light intensity exists, stretching from deep shadow to highly reflective surfaces in direct sunlight. It is now important to monitor the movements of the astronaut in deep shadow, so an entirely different coding scheme is appropriate. The results shown in figure 14 were obtained with the locally adaptive coding method based on the IDS model of retinal processing described in Section 2 above. Again, it should be possible to transmit only the data required for primal sketches at very high data compressions, or more complete data that permits the recovery of reflectance representations.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

ORIGINAL PAGE IS
OF POOR QUALITY

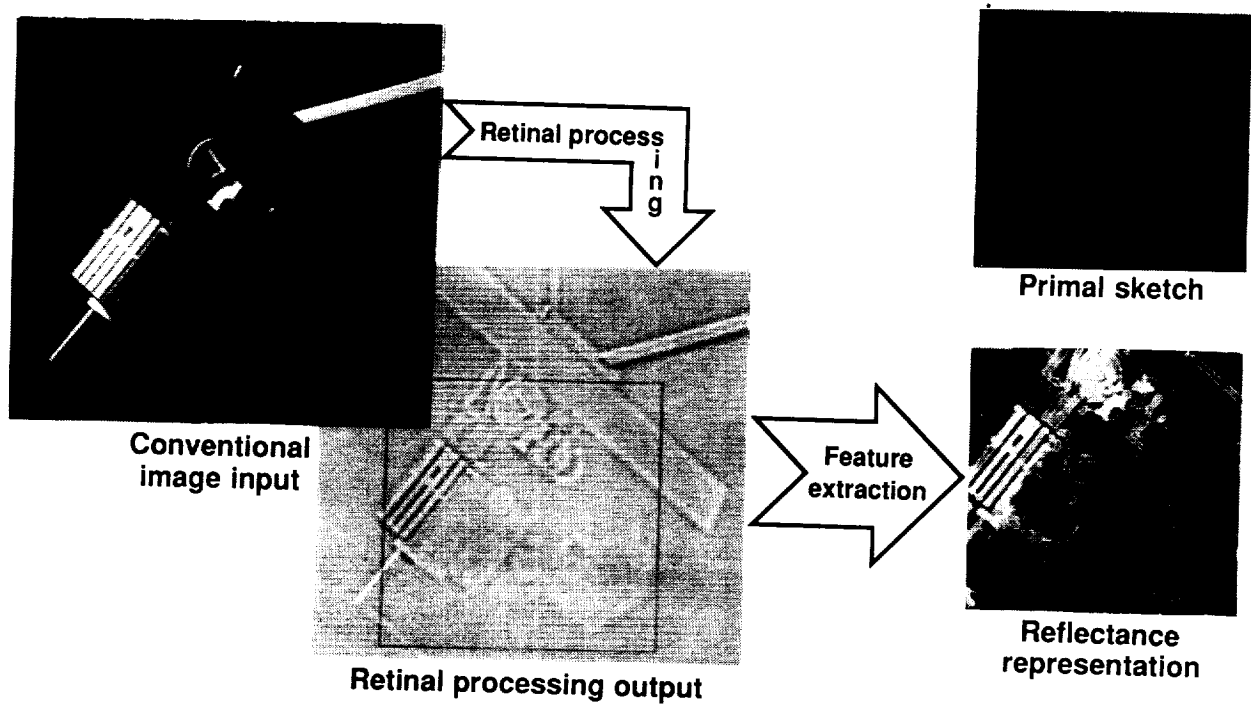


Figure 14. Feature extraction under conditions of extreme variations of light intensity, stretching from deep shadow to highly reflective surfaces in direct sunlight.

ORIGINAL PAGE IS
OF POOR QUALITY